DOCUMENT RESUME

ED 249 254 TM 840 596

AUTHOR Gullickson, Arlen R.
TITLE Matching Teacher Training with Teacher Needs in Testing.
PUB DATE 5 Mar 84
NOTE 35p.; Paper presented at the Annual Meeting of the American Educational Research Association (68th, New Orleans, LA, April 23-27, 1984). This study was supported in part by the University's General Research Fund.
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Educational Testing; Elementary School Teachers; Higher Education; Measurement; *Preservice Teacher Education; Secondary School Teachers; *Student Evaluation; Surveys; *Teacher Attitudes; *Teacher Education Curriculum; Teacher Educators

ABSTRACT

This study matched teacher and professor perspectives of preservice educational measurement courses. Twenty-eight professors from different colleges in seven states and 377 teachers from elementary and secondary schools in one midwestern state responded via mailed questionnaire in which: professors were asked to assess the emphasis they give to topics in preservice educational measurement courses, and precollege teachers were asked to state the emphasis they believe should be given to the topics. Results of the survey show that both groups gave high ratings to items on preparation of exams and low ratings related to the legal aspects of educational measurement. In other categories, their priorities differed. Teacher priorities were testing activities (the practical concerns of testing), nontest evaluative activities, and application of test results. Professor priorities were statistics, standardized test issues, and general use of exam results. Implications of these results are discussed. (Author/BW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Revised March 5, 1984

ED249254

Matching Teacher Training with Teacher Needs in Testing

Arlen R. Gullickson

University of South Dakota

This study was supported in part by the University's General Research Fund

Presented at the American Educational Research Association Conference in New Orleans April 26, 1984

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

A.R. Gullickson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TM 810 596

Running Head: Matching Teacher Training

Abstract

This study matched teacher and professor perspectives of preservice educational measurement courses. Twenty-eight professors from different colleges in seven states and 377 teachers from elementary and secondary schools in one midwestern state responded via mailed questionnaire in which: professors were asked to assess the emphasis they give to topics in preservice educational measurement courses, and precollege teachers were asked to state the emphasis they believe should be given to the topics. Results of the survey show that professor and teacher priorities match on one topic area, preparation of exams. In other areas, (statistics, standardized tests, nontest evaluation activities, and use of test results) their prioritites differed substantially. Implications of those results are discussed.

Matching Teacher Training with Teacher Needs in Testing

Conant (1963) sparked debate over the preparation of teachers, and in particular raised questions as to what should be teachers' preparation in educational measurements. Discussion ensued, and in a 1964 National Council on Measurement in Education (NCME) symposium Mayo (1964) reported measurement experts' expectations as to what teachers should be taught. Subsequent surveys (Goslin, 1967; Mayo, 1967; Roeder, 1972) established that teacher preparation in educational measurements did not meet these expectations. Stimulated by these findings, many colleges and universities introduced preservice educational measurement courses.

Once implemented, questions began to arise regarding the content of these courses. As early as 1972 Farr and Griffin argued that too much emphasis was being placed on standardized tests. More recently, research findings have raised questions about the course in general, and specific content issues as well. Pertinent research findings include: teachers report college courses do not properly prepare them in the use of instructional testing (Gullickson, in press); too little emphasis is placed on performance based

assessment (Stiggins & Bridgeford); teachers do not apply the statistical concepts taught in educational measurement courses (Gullickson & Ellwein, 1983); and teachers make only minor use of standardized tests for instructional purposes (Gullickson, 1983).

None of these studies expressly focused on what topics are covered in educational measurement courses and what emphasis should be given to the various topics. Thus, though the previous research suggests dissatisfaction with collegiate instruction efforts, it does not provide specific direction for improving the match between teacher needs and instructional content and emphasis. This study focused on that issue through a survey of both precollege teachers and college professors.

The objective was to identify teacher and professor priorities for these courses and, through contrast of their respective priorities, identify changes which would lead to a match between perceived teacher needs and measurement instruction. To do this both groups were presented with a comprehensive list of topics applicable to a preservice course in educational measurements and evaluation. Professors were asked to rate what emphasis they give to the topics; teachers

were asked to rate the emphasis they believe should be given to the topics.

Method

Subjects were elementary and secondary teachers from a midwestern state, and teachers of preservice educational measurement courses (hereafter called professors) from that state and the six contiguous states. Four hundred fifty teachers were randomly sampled from the state directory file of teachers in public and private schools. Previous research (Gullickson, 1983) with this population has shown that virtually all these teachers have a baccalaureate degree, the large majority received their training either in the state or in a contiguous state, and over 90% have had at least one course in educational measurements.

To obtain a sample of professors one-third (32) of the colleges in the seven state region were systematically sampled from the population of colleges and universities which give baccalaureate degrees in elementary and secondary education. Names of individual professors were obtained from the respective institutions through telephone calls to departments and individuals at each institution.

Procedures

Survey questionnaires, each with a cover letter which described the intended research and solicited the teacher's response, were mailed to 450 teachers. Subsequently, at one and one-half week intervals, nonrespondent teachers were mailed a reminder post card, a second letter with a second copy of the survey form, and finally a second reminder post card.

The designated professor at each institution was first called to solicit his/her participation in the survey, and then each was mailed a survey form. Non-respondent professors were subsequently sent a follow up reminder letter.

Professors were asked to respond to a variety of questions regarding their educational background, characteristics of the preservice educational measurements course, and content of the preservice measurement course. Teachers were asked only their teaching experience, their grade and curricular areas of instruction, and the parallel items to the professor's content of preservice measurement course questions.

This paper deals directly with the 67 course content items which the two questionnaires had in

common. The items (listed in Table 1) provide a comprehensive list of topics common to most undergraduate courses. The list was developed through analysis of textbooks (e.g., Hopkins & Stanley, 1981), and included topics identified by Rudman et. al. (1980) as relevant for measurement and evaluation courses. Professors from three nearby colleges reviewed the list for completeness. The items were blocked into eight categories: general assessment information, preparing examinations, administering and scoring tests, employing other evaluative devices, computing and interpreting statistical data, using test results for planning (formative evaluation) purposes, using test results for summative evaluation purposes, and testing and the law - legal challenges to test practices.

Teachers were asked to indicate the relative emphasis they believe <u>should</u> be placed on each of the topic areas, and professors were asked to respond regarding the emphasis they give to the topic in their instruction of the preservice measurement course. Both teachers and professors were asked to rate the items on a six point basis where 0 = no emphasis, i.e., not addressed; 1 = very slight emphasis; 2 = slightly greater emphasis; etc., up to 5 = very great emphasis.

Three hundred seventy-seven (84%) of the teachers

and twenty-eight (88%) of the professors responded. Among the professors who responded, two indicated that their college offered no distinct measurement course, and consequently they did not complete that section of the questionnaire. Therefore a total of 26 colleges are represented by the professor ratings.

Results

Teacher and Professor responses were summarized to provide item means and standard deviations for each group (Table 1). These item means were then rank ordered for each group. Finally the item mean differences between teacher and professor ratings were calculated and rank ordered. The top twenty ranked items from each list were then displayed in Tables 2-4. Tables 2 and 3 show the respective teacher and professor priorities, while Table 4 depicts the major discrepancies between the two groups.

Insert Table 1 about here

As can be seen in Table 1, teachers tended to rate items slightly more highly and were slightly more homogeneous in their average ratings. Both groups gave high rating to items in the category of Preparing Exams

and low ratings to items related to the legal aspects of educational measurements. In other categories their priorities differed.

Teacher priorities (Table 2), break naturally into three categories and are grouped accordingly as: testing activities (nine items), nontest evaluative activities (four items), and application of test results (seven items). The first category, testing activities, depicts practical concerns of testing. These items refer to content definition, selecting and/or writing test items, the administration of the test, and scoring test results.

Insert Table 2 About Here

Among the nontest activities class discussion and observing work habits stand out as particularly important because: a) teachers rated these items very highly; the average for each was above 4.0, and b) of all the 67 items teacher and professor ratings were most discrepant for these two. Clearly teachers value these practices much more highly than do professors.

Application of test results directly relates to decisions which a teacher could/would make based upon test results. Three of these items (items ranked 11, 14,

19) are instruction oriented; four (items ranked 3, 6, 7, 17) are student oriented. In composite they indicate teacher interest in planning coursework and follow-up based on test results to improve student learning. In particular teachers appear to prioritize use of tests to get at "outlier characteristics" both for individual students (e.g., strengths/weaknesses) and for student groups within the classroom (e.g., gifted, slow, and underachieving learners).

Professor priorities (Table 3) divided into four categories: statistics (six items), preparation of exams(eight items), standardized test issues (four items), and general use of exam results (two items). Of the four, professors give greatest emphasis to the instruction of statistics and preparation of exams. Those two comprise all of the top ten rated items.

Insert Table 3 about here

Where teachers included no statistical items in their top twenty, professors included five in their top ten. Essentially every major statistical topic presented in standard measurement and evaluation textbooks is represented on the professors' list as

having a high priority.

Like teachers, professors prioritize the preparation of exams. However, professors give priority to only the front end of the testing process, i.e., preparation. None of the priority items refer to the administration or scoring of tests.

Two items in this category deserve special note. These two items, ranked 8 and 9, speak respectively to the comparative merit of items (taxonomy levels), and the comparative merits of testing strategies (objective vs. essay tests). Neither item was among the top 20 for teachers. In fact the first of the two, which speaks directly to skill and taxonomy levels, is one of the items for which teacher and professor ratings are most discrepant.

Professors prioritize each of the three major standardized test issues: norms, reliability, and validity. These testing issues, when combined with the topic of statistics (which is central to the interpretation of standardized tests), result in a substantial emphasis on standardized testing concerns.

The professors' application of test results category contains substantially fewer items than did the teachers', and is substantially different in content. Where teachers prioritized several specific

applications, professors included none. Instead professors included only items which refer in general terms to the use of test results for formative and summative evaluation purposes. This suggests an interest in the general area but either specific topics relevant to the professors were not included on the instrument or professors focus on providing a general overview rather than application of test results to specific concerns.

The major discrepancies, Table 4, complement the two priority lists. Here many of the individual items are different from those included in Tables 2 and 3 but the areas of major discrepancies are quite consistent with the differential priority patterns noted above.

Insert Table 4 about here

Examination of the discrepant items reveals four major categories: standardized tests, statistics, nontest evaluative activities, and use of test results. All four represent priority categories for either professors or teachers. Two remaining items, ranked 12 and 16 do not clearly represent a single category. One regarding taxonomy levels was noted previously, as one

of the professor priorities within the preparing exams category. The other appears to suggest teachers place a higher priority on learning about resource information.

The categories with greatest definition and direction were nontest evaluative activities and statistics. Three of the four nontest evaluation items listed as priorities by teachers appear in this list. Similarly of the five statistical concepts which appear in this list, four were listed as priorities by professors. (The fifth, though it has a clear statistical connotation, was included in the questionnaire in the Preparing Examinations section.)

Both categories reemphasize the difference in teacher and professor priorities. Teachers value nontest activities but not statistics; for professors the opposite is true.

Use of test results emerged as the third major category of discrepancy. Only one of these items was also listed as a priority. But these items confirm what was evident in the priority tables. Teachers place a much higher priority on learning to use test results for specific decision situations.

Of the four discrepant items related to standardized tests, two are professor priorities. The

remaining two, though not teacher priorities, depict teacher preferences. One refers to the use of test committees, the other regards the recommendation of tests to administrative personnel. In general, these four items suggest a professor focus on the aspects of standardized tests which are necessary for determining which tests to use, and a teacher focus on practical issues which teachers face in the selection of standardized tests.

Discussion

Only for preparing examinations was there a clear match between instructional emphasis and perceived needs. Even there noteworthy differences emerged. Professors lean toward emphasis on fundamental and qualitative concepts (course objectives, comparative merits of tests, etc.). Teachers however seem to give higher priority to practical concerns: the preparation of test items, administration of tests, and scoring of tests. Both professors and teachers did prioritize use of test results, but professor priorities were focused at a very general level, and not on the more specific issues desired by teachers.

Teachers clearly do not share professor priorities for statistics; this is consistent with previous survey

findings Mayo (1967). In fact as already noted, despite its heavy emphasis in measurement courses, teachers do not use statistics. Teachers also give a lower priority to standardized test concerns. Again this is consistent with their minimal use of such tests for instructional purposes.

Less obvious but important is the discrepancy on issues of test validity. Teachers would choose to give less emphasis to test validity issues than do professors. This is evident in the questions which speak directly to validity and in other items which refer to the relative characteristics of tests and test items.

These striking differences between course emphases and teacher preferences probably result from the differing conceptual frameworks and contextual constraints surrounding teachers and professors. Most teachers have a very limited formal preparation in educational measurements and evaluation; but, they have a substantial faith in the validity of tests (Farr & Griffin, 1972). Also, because of the demands of their job, teachers extensively use tests and other evaluation techniques. Calberg (1980) estimates that approximately 32% of a teacher's time is vested in test related activities. They routinely prepare administer

and score their own exams, evaluate their students through the use of nontest techniques, and apply the results of their tests and evaluations to students and their teaching situations (Gullickson, 1983). Such conditions certainly would stimulate an interest in practical application concerns.

Professors, on the other hand, have a substantial background in educational measurements. (The group surveyed typically had from eight to 15 semester hours in educational measurement courses at the graduate level.) Typically such graduate courses emphasize topics such as standardized tests and statistics. Always, such courses place a heavy emphasis on test validity, and informal, nontest, approaches are frequently cited more for their validity problems than for their evaluation potential.

In addition professors routinely teach measurement courses under adverse conditions. For example professors typically: a) have three semester hours or less in which to teach educational measurements and evaluation, b) teach all preservice teachers in a single group, or teach preservice elementary teachers in one group and secondary teachers in another, and c) do not teach the course in the context of any

individual curriculum or even in the context of curricular content. Rather it is usually taught as a unique course focusing on facets of testing germane to all curricular areas, or it is taught as a part of an educational psychology course.

In such circumstances it is easy to see why professors give priority to statistics and standardized tests. They have a strong background in both areas. Plus statistical concepts are important to the analysis and interpretation of both standardized tests and teacher made tests. Further, while class discussion and other nontest evaluation activities typically suffer from validity concerns, standardized tests (particularly achievement tests) tend to be valid. In addition standardized testing concerns are readily generalized across curricular areas, and published standardized tests can be used to exemplify attributes which tests should and should not have.

While the respective conditions help to explain why the discrepancy exists, they are not a rationale for maintenance of such discrepancies. Previous research suggests that teachers, in practice, ignore measurement prescription that does not meet their needs; and clearly, present courses are not meeting their perceived needs.

The results suggest teachers see their needs as practical in nature; there is little in the teacher priorities which point to conceptual concerns. Thus, the problem professors face is to produce measurement courses which fit teacher perceived needs yet are conceptually sound. To meet those constraints it appears measurement concepts will need to framed within practical application situations.

Presently statistics and standardized tests retain a central role with implications drawn from them for teacher made tests and evaluations. Given the present findings, it seems appropriate to reverse the emphases. That is teacher made tests and other evaluative devices can be given central focus, and standardized tests and statistics can be included as supporting tools.

Such a shift would have substantial implications for instruction of validity concepts. Validity has routinely been addressed from the context of standardized tests, and the majority of research on this topic is related to standardized test use. To date measurement experts have simply assumed standardized tests and teacher made evaluation techniques share the same validity concerns. There is therefore a need to gain more information about teacher

evaluation practices: to identify typical errors, to identify strategies to deal with these errors within the limits and resources of the teaching situation, and to determine the proper emphasis to give the topic in preservice teacher preparation courses.

Finally, it seems appropriate to change the context within which the educational measurements course is offered. Actually, the question of course format has been an issue from the outset. Findley (1964) argued, in the 1964 NCME symposium, for instruction of educational measurements not as a separate course but as a part of the methods course in the curricular content areas. If the measurement content were included in those courses, instructional groups would be more homogeneous and measurement concepts would be taught within the context of content areas to be evaluated.

Unfortunately such action would be likely to "create" different problems. Personal experience suggests that few professors, who teach teaching methods for the various content areas, are skilled in measurement and evaluation methodology. For example, at the local university a check with the computer center revealed that none of the "methods professors" ever ask for an item analysis of their tests (this

service is provided free of charge to any professor who requests it).

Given these circumstances, it may be appropriate to teach educational measurement courses in tandem with the methods courses so that the general ideas gained from the measurement course can be given specific application in the methods course. Or alternatively, professors of educational measurements and evaluation could provide inservice to methods instructors who could then carry out preservice measurement/evaluation instruction. Both actions require greater coordination of instruction than is typical at the college level, but both seem likely to yield a better match with precollege teacher needs.

Uniformly these suggestion have called for changes at the college level to make the courses coincide with teacher perceived needs. This does not suggest that teachers properly perceive their own needs. Almost certainly other and different steps will have to be taken before educational measurement courses truly match teachers needs. Yet, these steps seem to be an appropriate beginning.

References

Conant, J. B. (1963). The Education of American Teachers. New York: McGraw Hill.

Farr, R., & Griffin, M. (1973). Measurement Gaps in Teacher Education. Journal of Research and Development in Education. 7(1), 19-28.

Findley, W. G. (1964). When and how to teach educational measurements. Journal of Educational Measurement. 1(1), 87-89.

Gullickson, A. R., & Ellwein, M. C. (1983). The goodness-of-fit between prescription and practice: Post hoc analysis of teacher-made tests. Unpublished manuscript, University of South Dakota, Vermillion.

Gullickson, A. R. (1983). The practice of testing in elementary and secondary schools. Paper presented at the 1982 Rural Education Conference at Kansas State University Manhattan Kansas. (ERIC Document Reproduction Service No. ED 229 391)

Gullickson, A. R. (in press). Teacher perspectives of their instructional use of tests. Journal of Educational Research.

Goslin, D. A. (1967). Teachers and testing. New York: Russell Sage Foundation.

Hopkins, K. D., & Stanley, J. C. (1981). Educational and psychological measurement and evaluation

(6th ed.). Englewood Cliffs, NJ: Prentice Hall.

Mayo, S. T. (1964). What experts think teachers ought to know about educational measurement. Journal of Educational Measurement. 1(1), 79-86.

Mayo, S. T. (1967). Pre-Service Preparation of Teachers in Educational Measurement (Contract No. OE 4-10-011). Chicago, IL: Loyola University.

Roeder, H. H. (1972) Are today's teachers prepared to use tests? Peabody Journal of Education. 49(3), 239-240.

Rudman, H. C., Kelly, J. L., Wanous, D. S., Mehrens, W. A., Clark, C. M., & Porter, A. C. (1980). Integrating assessment with instruction: A review (1922-1980). Institute for Research on Teaching, Michigan State University, College of Education.

SAS Institute, Inc. (1979). SAS user's guide, 1979 edition. Raleigh, North Carolina: Author.

Stiggins, R. J., & Bridgeford, N. J. (1982). Final Research Report on the Role, Nature and Quality of Classroom Performance Assessment (Contract No. 400-80-0105). Portland OR: Northwest Regional Educational Laboratory, Center for Performance Assessment.

Table 1

Teacher and Professor Means and Standard Deviations for Individual Questionnaire Items

Teachers		Professors		Item Description
M	S	M	S	
3.36	1.10	3.10	1.07	1. General Assessment Information
3.20	1.13	2.95	1.23	A. Sources of aid in interpreting and using assessment information
3.19	1.21	2.50	1.25	1. interpretative manuals and devices which accompany published tests
2.96	1.17	2.08	1.38	2. orientation and instructional information available to teachers through publishers and major organizations, e.g., Psychological Corporation and the National Educational Association
3.18	1.26	3.72	0.89	B. Selection and use of standardized and publisher prepared tests
3.50	1.24	3.13	1.32	1. identifying tests suitable for particular age or grade groups
3.67	1.14	3.17	1.40	2. determining appropriateness of test content for specific classes
3.56	1.24	3.50	1.29	3. evaluating tests in terms of reliability and validity

(table continues)

Matching teacher 24

Teachers		Professors		Item Description
M	S	M	S	
2.77	1.32	1.83	1.40	4. recommending tests to administrative personnel
2.93	1.15	3.67	1.43	5. test norms and interpretation based upon norms
3.44	1.24	2.88	1.73	6. critical review of individual tests
4.10	1.06	4.00	1.33	2. Preparing Examinations
3.53	1.07	3.38	1.40	A. General development concerns
2.50	1.22	1.71	1.46	1. working on test committees
3.69	1.17	3.71	1.68	2. defining course objectives
2.89	1.16	3.67	1.34	3. norm referenced vs. criterion referenced tests
3.58	1.11	3.79	1.25	4. objective vs. essay tests
2.85	1.18	3.79	1.47	5. defining skills and taxonomy levels (e.g., Bloom's Taxonomy)
3.90	1.08	3.96	1.37	6. writing test items
3.53	1.17	2.63	1.74	7. determining scoring weights
3.82	1.03	3.74	1.59	B. Item selection and construction
3.80	1.08	3.54	1.53	1. writing subjective test items, e.g., essay and short answer items
3.86	1.01	3.75	1.45	2. writing objective items, e.g., multiple choice and matching items

(table continues)

Matching teacher 25

Teachers		Professors		Item Description
M	S	M	S	
3.12	1.15	3.25	1.39	3. conducting an item analysis
3.67	1.11	3.37	1.21	3. Administering and Scoring Tests
3.48	1.12	3.04	1.26	a. Establishing test taking routines
3.37	1.11	2.91	1.36	b. Administering informal tests
3.13	1.25	2.87	1.46	c. Administering standardized tests
3.63	1.11	3.26	1.36	d. Scoring tests
3.14	1.24	3.52	1.34	e. Transforming raw scores
3.41	1.23	2.70	1.52	f. Recording test results
3.60	1.01	3.40	0.94	4. Employing Other Evaluative Devices
3.14	1.11	3.22	1.13	a. Use of check lists and rating scales to evaluate procedures and products
3.46	1.23	2.22	1.62	b. Interviewing pupils and parents
4.01	1.01	2.43	1.50	c. Observing work habits
4.06	0.95	2.39	1.64	d. Class discussions
3.68	1.17	2.17	1.64	e. Interpersonal relationships
2.67	1.16	2.22	1.59	f. Administering sociometric tests
2.80	1.26	2.18	1.37	g. Writing anecdotal reports
2.81	1.12	3.90	0.83	5. Computing and Interpreting Statistical Data
2.58	1.06	4.00	1.13	a. Measures of central tendency and variability

(table continues)

Matching teacher 26

Teachers		Professors		Item Description
M	S	M	S	
3.10	1.13	3.83	1.23	b. Percentages and percentiles
2.92	1.12	4.00	1.00	c. Standard scores and the normal distribution
2.87	1.22	3.17	1.19	d. Reading graphs
2.57	1.18	3.83	1.19	e. Correlations and reliability coefficents
2.66	1.21	3.04	1.36	f. Reading statistical reports of test results
3.76	1.03	3.65	0.67	6. Using Test Results for Planning (Formative Evaluation) Purposes
4.05	0.99	3.39	0.94	a. Interpreting test profiles to identify pupil strengths and weaknesses
3.63	1.19	3.04	1.30	b. Pretesting to determine required instructional emphases
3.11	1.24	2.22	1.31	c. Organizing homogeneous groups on the basis of test performance
3.68	1.06	3.04	1.40	d. Using test data to guide remediation
3.48	1.00	2.78	1.24	e. Examining test data to anticipate level of class performance
3.40	1.27	3.13	1.25	f. Using test scores to evaluate teaching

(table continues)

Teachers		Professors		Item Description
M	S	M	S	
3.47	1.03	3.55	1.00	7. Using Test Results for Summative Evaluation Purposes
3.64	1.00	2.87	1.29	a. Recommending counseling or remediation
3.60	1.16	3.26	1.45	b. Assigning final grades
3.89	1.06	3.17	1.37	c. Identifying gifted pupils or slow learners
3.86	1.00	3.39	1.20	d. Identifying under-achievers
3.51	1.18	2.61	1.23	e. Recommending promotion, acceleration, or retardation
2.73	1.37	1.48	1.34	f. Recommending employment
3.14	1.31	1.96	1.55	g. Recommending higher education
2.74	1.36	2.09	1.31	8. Testing and the Law. Legal Challenges to Test Practices
2.48	1.30	1.91	1.34	a. Court cases and major considerations in those court cases
2.52	1.33	2.00	1.45	b. Major rulings of courts
2.83	1.32	2.29	1.31	c. Legal restrictions
2.76	1.37	2.36	1.29	1. on use of standardized tests
2.90	1.37	2.41	1.40	2. on use of teacher prepared tests

Note. The rating scale ranged from 0 = no emphasis to 5 = very great emphasis.

Table 2

Teacher and Professor (P) Item Means for the 20 Items Teachers Rated Most Highly

Teacher Rank	Teacher Mean	P Mean	Item Descriptions
			Preparing Examinations
1	4.10	2.39	Preparing Examinations
5	3.90	3.96	Writing test items.
8	3.86	3.75	writing objective items, e.g., multiple choice and matching items.
9	3.82	3.74	Item selection and construction
10	3.80	3.54	Writing Subjective test items, e.g., essay and short answer items.
12	3.69	3.71	Defining course objectives
15	3.67	3.37	Administering and scoring tests.
16	3.67	3.17	Determining appropriateness of test content for specific classes.
18	3.63	3.26	Scoring tests.
			Nontest Evaluation Activities
2	4.06	2.39	Class discussion
4	4.01	2.43	Observing working habits.

(table continues)

Teacher Rank	Teacher Mean	Prof. Mean	Item Descriptions
13	3.68	2.17	Interpersonal relationships.
20	3.60	3.40	Employing other evaluative devices.
			Application of Test Results
3	4.05	3.39	Interpreting test profiles to identify pupil strengths and weaknesses.
6	3.89	3.17	Identifying gifted pupils or slow learners.
7	3.86	3.39	Identifying under-achievers.
11	3.76	3.65	Using test results for planning (formative evaluation) purposes.
14	3.68	3.04	Using test data to guide remediation.
17	3.64	2.87	Recommending counseling or remediation.
19	3.63	3.04	Pretesting to determine required. instructional emphases.

Table 3

Teacher (T) and Professor (P) Item Means for the 20 Items Professors Rated Most Highly

P Rank	T Mean	P Mean	Item Description
			Statistics
1	2.92	4.00	Standard scores and the normal distribution
3	2.58	4.00	Measures of central tendency and variability.
5	2.81	3.90	Computing and interpreting statistical data.
6	2.57	3.83	Correlations and reliability coefficients.
7	3.10	3.83	Percentages and percentiles.
19	3.14	3.52	Transforming raw scores.
			Preparation of Examinations
2	4.10	4.00	Preparing examinations
4	3.90	3.96	Writing test items.
8	2.85	3.79	Defining skills and taxonomy levels (e.g., Bloom's taxonomy).
9	3.58	3.75	Objective vs. essay tests.
10	3.86	3.74	Writing objective items, e.g., multiple choice and matching items.

(table continues)

Matching teacher 31

P Rank	T Mean	P Mean	Item Description
11	3.82	3.74	Item selection and construction.
13	3.69	3.71	Defining course objectives.
18	3.80	3.54	Writing subjective test items, e.g., essay and short answer items.
			Standardized Test Issues
12	3.18	3.72	Selection and use of standardized and publisher prepared tests.
14	2.89	3.67	Norm-referenced vs. criterion-referenced tests.
15	2.93	3.67	Test norms and interpretation based upon norms.
20	3.56	3.50	Evaluating tests in terms of reliability and validity.
			General Use of Exam Results
16	3.76	3.65	Using test results for planning (formative evaluation) purposes.
17	3.47	3.55	Using test results for summative evaluative purposes

Table 4

Major Mean Discrepancies (MD) Between Teacher and Professor Mean Ratings

Rank	MD	Item Description
		Nontest Evaluation Activities
1	1.67	Class Discussions.
2	1.58	Observing work habits.
3	1.51	Interpersonal Relationships.
7	1.24	Interviewing pupils and parents.
		Statistics
4	-1.42	Measures of central tendency and variability.
5	-1.26	Correlations and reliability coefficients.
9	-1.09	Computing and interpreting statistical data.
10	-1.08	Standard scores and the normal distribution.
14	0.90	Determining scoring weights.

(table continues)

Matching teacher 33

Rank	[illegible]	Item Description
		Standardized Tests
17	0.79	Working on test committees.
18	-0.78	Norm referenced vs. Criterion referenced tests.
		Use of Test Results
20	-0.74	Test norms and interpretation based upon norms.
6	1.25	Recommending employment
8	1.18	Recommending higher education.
11	0.94	Recommending tests to administrative personnel.
13	0.90	Recommending promotion, acceleration, or retardation.
15	0.90	Organizing homogeneous groups on the basis of test performance.
19	0.77	Recommending counseling or remediation.

(table continues)

Matching teacher 34

Rank	MD	Item Description
		Other
12	-0.94	Defining skills and taxonomy levels (e.g., Bloom's taxonomy).
16	0.88	Orientation and instructional information.

Note. If MD is positive, teachers had the larger mean rating.